



Queen's Economics Department Working Paper No. 947

WAGE AND TEST SCORE DISPERSION SOME INTERNATIONAL EVIDENCE

Kelly Bedard
McMaster University

Christopher Ferrall
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

03-1997

WAGE AND TEST SCORE DISPERSION: SOME INTERNATIONAL EVIDENCE

Kelly Bedard
Department of Economics
McMaster University
Hamilton, Ontario
bedardk@mcmill.cis.mcmaster.ca

Christopher Ferrall
Department of Economics
Queen's University
Kingston, Ontario
ferrallc@post.queensu.ca

March, 1997

Abstract

We study fifty observations on wage distributions across eleven countries and two age cohorts defined by international mathematics tests given to thirteen-year-olds in 1962 and 1982. We find that wage dispersion later in life is never greater than test score dispersion. In particular, Lorenz curves for a cohort's wages always lie above or on top of the cohort's test score Lorenz curve. Wage dispersion, as summarized by Gini coefficients, is significantly related to test score dispersion and union density in the country. A general fall in test score dispersion between 1962 and 1982 appears to be reflected in reduced wage dispersion. For three countries with available data (the U.S., the U.K., and Japan), we find evidence of skill-biased changes in wage dispersion between the early 1970s and the late 1980s.

JEL Classification: I2, J3

Keywords: Mathematics Test Scores, Wage Distributions, Lorenz curves

Bedard thanks the Canadian International Labour Network (CILN) for financial support. Ferrall acknowledges research support from the Social Sciences and Humanities Research Council of Canada. Two preliminary versions of this paper were circulated under the title "Test Scores and Wage Inequality: Some International Evidence." For comments and suggestions on those versions that helped us converge to the current paper, we thank Charles Beach, Russell Davidson, Allan Gregory, Peter Kuhn, Craig Riddell, and seminar participants at Lakehead, Brock, and the CILN Conference held in Burlington, Ontario in 1996.

I. Introduction

Recent studies of wage inequality have documented large differences in inequality across countries (e.g. Gottschalk and Joyce 1992), a trend in some countries towards increased inequality from the 1970s to the 1980s (e.g. Davis 1992), and the importance of wage-setting institutions in explaining international differences (Blau and Kahn 1996). While institutions such as unions and the minimum wage appear to partially explain international differences in wages, there is perhaps no greater difference between adult populations in different countries than the school systems they went through as children. We find support for this hypothesis by comparing the Lorenz curve of test scores for mathematics exams given to thirteen-year-olds and the Lorenz curve of wages later in life.

Our analysis is based on fifty observations across eleven countries, two birth cohorts, and several ages at which wages can be measured for the cohorts. The test score data come from the First (1962) and Second (1982) International Mathematics Examinations (IME). The exams define two cohorts for which we calculate dispersion in annual wages later in life using the Luxembourg Income Study (LIS) database and other sources. These data sources, as with all other sources we know of, do not provide test scores and wages later in life for individuals in several countries. Despite this limitation, the Lorenz curves provide novel evidence about the link between school systems and labor market outcomes.

We find that the Lorenz curve for wages always dominates the Lorenz curve for test scores. That is, except for some borderline cases, the test scores of thirteen-year-olds are never less disperse than the distribution of wages later in life. Test scores tend to be significantly more unequal than wages for the first IME cohort and much closer together (and in some cases nearly identical) for the second cohort. However, the gap between the Lorenz curves tends to be largest at low percentiles for both cohorts, which might suggest that wage equalization policies would explain the gap. We find that union density does help explain the gap between wage and test score Lorenz curves, as measured by Gini coefficients.

Can dispersion in test scores at age thirteen be explained by measurable characteristics of primary schools? We find that pupil-teacher ratios are *negatively* related to test score dispersion, after controlling for IME cohort. Countries in which primary school teachers have more students not only have lower dispersion, but also higher median scores. It is not entirely clear whether the differences in test scores between the birth cohorts are due to changes in the testing instrument or whether they are due to a general change in educational practices that occurred between the 1950s and 1970s, when our cohorts were in primary school. However, we do find a steady relationship between the wage and test score Gini coefficients across the two cohorts, suggesting that the drop in test score dispersion is not an artifact of changes in the IME test, but instead reflects changes in the distribution of skills across the two cohorts. On the other hand, for three countries we are able to compare Lorenz curves for the two cohorts at roughly the same age, and we find that wages are more disperse for the young IME cohort than would have been predicted from their test scores and the experiences of the older cohort.

II. Preliminaries

Mathematics Tests

In making international comparisons, mathematics is perhaps the best subject to focus on. In most countries, twelve to fifteen percent of class time is allocated to the study of mathematics (Robitaille, 1990). Mathematics exams therefore measure the performance of a major component of the curriculum. Further, many other skills used in the workplace are based on math skills. Taubman and Wales (1974) and Paglin and Rufolo (1990) find that math scores are the best academic predictor of future wages in the United States. The only other subject of similar importance to mathematics is language, which is inherently much more difficult to assess internationally.

The First and Second International Mathematics Examinations (IME) were conducted

in 1962 and 1982 by the International Association for the Evaluation of Educational Achievement (IEEA). Twelve countries participated in the first IME (cohort I), and twenty-two countries participated in the second (cohort II). While the first IME test instrument and sampling procedure were initially criticized (Freudenthal, 1975), more recent studies have supported its validity (Keeves 1988). A committee of national representatives used curricula outlines and test construction recommendations submitted by participating countries to develop a test guideline that included the test format, topics, and prototype questions. The committee then invited participating countries to submit section-specific questions that complied with the guidelines. Using new and submitted questions the committee of representatives produced a preliminary exam which was circulated to national testing centers for preliminary testing and feedback. After reviewing the results, the final test instruments were agreed upon.

To ensure that the samples would be representative, each country was stratified by geographic region and schools were randomly selected from within each region.¹ Students were selected to take exams at two points in their schooling when they could be identified consistently across countries. Namely, the IEA gave an exam to all thirteen-year-olds and another exam to students enrolled in mathematics during their pre-university year. Since differences in high school curricula and participation rates make it difficult to define comparable samples of older students, we focus on the exams given to thirteen-year-olds. All participating countries had 100% school participation at this age in both 1962 and 1982.²

Table 1 lists the countries that participated in the first (cohort I) and second (cohort II) IME that are used in our study. Of the twelve countries participating in the first IME, we combine England and Scotland, and we exclude Israel because the high rate of immigration to Israel makes it dubious to link test scores in the 1960s to wages in the 1980s. Of the

¹ West Germany is the only exception. West Germany made participation voluntary and then randomly selected schools from the set of volunteers.

² In countries where streaming of students had already occurred representative samples were constructed.

twenty-two countries participating in the second IME, we use all countries for which we could obtain the necessary wage data. The Canadian provinces of British Columbia and Ontario joined the second IME and we were able to gather wage data separately for them. The only countries that fall out from the first IME are West Germany and Australia.

Wages Later in Life

Table 1 also lists all wage sample information for both IME cohorts, including data sources, sample sizes and sample years.³ The data for all but two countries were drawn from the the Luxembourg Income Study (LIS) database. To control for factors other than ability that determine (annual) wages, we selected full time male workers who are not self-employed and who would have been approximately 13 years old at the time of the IME exam.⁴ We use the sampling criteria to control for many of the individual characteristics that Gottschalk and Joyce (1992) and Blau and Kahn (1996) controlled for using variables in wage regressions. While it is also possible to control for years of education and occupation, we purposely do not do this because we are interested in the net link between early test scores and wages later in life. As Card and Krueger (1992) suggest, differences across school systems may ultimately encourage different educational attainment and different occupational choices. Whether math skills feed directly into wages or instead lead to different educational paths that then feed into wages is not a question we can address.

Lorenz Curves and Gini Coefficients

Since wages are measured in local currencies and math skills are measured in the number of correct answers on an exam, it is clear that these variables must be normalized before they can be compared to each other or internationally. Because we have the full distribution of wages and test scores, the Lorenz curve is perhaps the best way to compare and contrast

³ To our knowledge, this is the first attempt to link the IME with wages later in life.

⁴ We cannot separate male and female math test scores, so the test score distributions include all students.

dispersion in them.⁵ Let Y be a random variable with density $f(y)$, inverse distribution function $F^{-1}(z)$, and mean μ . The percentiles of Y are given by $y(z) = F^{-1}(z)$ and the Lorenz curve for Y is

$$L(z) \equiv \frac{\int_0^z y(z)f(y(z))dz}{\mu}.$$

Let $L_s(z)$ and $L_w(z)$ be the Lorenz curve for test scores and wages, respectively. $L_s(z)$ is the proportion of right answers earned by the lowest 100z% scoring students in the cohort, while $L_w(z)$ is the proportion of wage income earned by the lowest paid 100z% people in the same cohort later in life. (For several countries we can measure $L_w(z)$ at more than one age.) As is well known, the Lorenz curve for the uniform distribution is simply $L(z) = z$, and the Gini coefficient is defined as its total deviation from uniformity:

$$G \equiv 2 \int_0^1 (z - L(z))dz.$$

Let G_s and G_w denote the test score and wage Gini coefficients, respectively.⁶

Why might $L_s(z)$ and $L_w(z)$ be related?

If a person's wage-earning power is related to their math skills, then $L_s(z)$ should be related to $L_w(z)$. For example, suppose in the most basic case that an individual's productivity is proportional to their test score:

$$w(s) = \alpha s.$$

Since the Lorenz curve is insensitive to proportional transformations, $L_w(z) = L_s(z)$ for all z and $G_s = G_w$. This goes beyond saying that wages tend to be ranked by math skills, which as a hypothesis does not determine any relationship between $L_w(z)$ and $L_s(z)$, since the Lorenz curve internally rank the data anyway.

Even if the IME tests skills used in the labor market, there are at least four reasons why $L_w(z)$ should not equal $L_s(z)$: centralized wage-setting, signaling of ability, pursuit of

⁵ Some percentiles were linearly interpolated due to the grouping in the IME data and the Japanese wage data.

⁶ In this paper the reported Gini coefficients are the usual definition divided by 2.

comparative advantage, and skill-biased education attainment. Wage-setting institutions include factors such as unions and high minimum wages. Institutions such as these that are progressive in wages (although not necessarily in employment opportunities) would tend to equalize wages and would reduce the link between skills and wages. For example, wages might be related to scores by:

$$w(s) = \begin{cases} \bar{w} & \text{if } s < \bar{s} \\ \bar{w} + \alpha(s - \bar{s}) & \text{if } s \geq \bar{s} \end{cases}$$

Here \bar{w} might be the wage paid in the minimum wage or unionized sector, which attracts workers with low math skills. This type of wage function could be the result of signaling as well, as low-skill workers may choose not to signal productivity and hence receive a pooling wage. In either event, the wage distribution is compressed at the bottom end relative to the test score distribution, leading to $L_w(z) > L_s(z)$ at low percentiles, and a narrowing gap as z increases. In turn, the Gini coefficients will differ, with $G_w < G_s$.

Students with low math skills may also have other skills not measured by a math test, but that nonetheless have value in the labor market. For example, literacy and physical skills may not be perfectly correlated with mathematical skills. As with unions and minimum wages, the ability to pursue comparative advantage outlined by Roy (1951) will make wages less unequal than test scores. If l is the other (composite) skill in a simple two-sector Roy model, then the pursuit of comparative advantage leads to a wage function of the form:

$$w(s, l) = \begin{cases} \beta l & \text{if } s \leq \beta l / \alpha \\ \alpha s & \text{if } s > \beta l / \alpha \end{cases}$$

where β is the price of skill l . Again, the Lorenz curve for wages lies above that of math test scores, because the average poor math student is using l and earning a wage more than proportional to their math skill s . In turn, the Gini coefficient for wages would be greater than the coefficient for scores.

The effect of education and other human capital accumulated after age thirteen will also affect $L_w(z)$. It is unclear whether further schooling exaggerates or dampens skill dispersion.

If students can develop skills that they have comparative advantages in, then further schooling will compress the effective ability distribution. That is, poor performers in math may focus on mechanical and literacy skills and then choose occupations using those skills more intensely than mathematical skills. However, school systems have many rigidities and sorting mechanisms that do not necessarily encourage or allow ‘pure’ self-selection. Bedard (1997) studies the effect of academic streaming on occupational choice and wage distributions in West Germany.

On the other hand, if a student’s share of further school resources is tied to the students standing at age thirteen, then the school system will compound skill differences. Those with high test scores will accumulate skills at a faster rate than low scorers, and will go on to earn an even greater share of wages (as compared to their skill share at age thirteen). For example, the wage function might take the form

$$w(s) = \begin{cases} \alpha s & \text{if } s < \bar{s} \\ \alpha(1 + \tau)s & \text{if } s > \bar{s} \end{cases}$$

where $\tau > 0$ is the extra augmentation of skills that ‘star’ pupils receive after age thirteen. Here, we would find $L_w(z) \leq L_s(z)$ and wages would be more unequal than math skills. Perhaps comparative advantage dominates for low scoring students, while skill-augmentation dominates for high scoring students. In this case, the Lorenz curves may cross with $L_s(z) > L_w(z)$ for $z < z^*$ and $L_s(z) \leq L_w(z)$ for $z \geq z^*$. The magnitude of G_w and G_s would be ambiguous.

Finally, there is the distinct possibility that the IME test measures nothing meaningful or that the distribution of math skills at age thirteen is irrelevant to wages later in life. In that case, there would be no obvious relationship between $L_w(z)$ and $L_s(z)$ and no theoretical ranking of the magnitudes of G_w and G_s . Given this possibility, and the ambiguous effects of some factors outlined above, any consistent relationship across countries between wages and tests taken a quarter century earlier may be surprising.

III. Results

Wage and Test Score Dispersion Within IME Cohorts

Table 2 summarizes the test score and wage distributions derived from our data sources reported in Table 1. In both IME cohorts the United Kingdom had the most disperse, or unequal distribution of scores as measured by the Gini coefficient. Finland had the lowest amount of dispersion in the 1962 exam, and France had the lowest amount in 1982 (Japan follows quite closely). In terms of median scores, Sweden and Japan had the lowest and highest on both tests, although some countries switched rankings between exams.

There may be a tradeoff between a high level of scores and an unequal distribution of scores. That is, countries with low test score dispersion may be sacrificing the test score of the median student. This does not appear to be the case. Figure 1 shows that the Gini coefficient and the median score are negatively related. There is a sharp increase in median scores from the first to the second cohort which is accompanied by a decline in wage dispersion. A regression shows that much, but not all of the decline can be explained by a consistent tradeoff between median scores and wage dispersion:

$$\begin{aligned}
 G_s &= .2330 & - & .006 \times \text{Median} & - & .0177 \times \text{CohortII} \\
 (12.4) & & (3.95) & & (1.44) \\
 F_{2,17} &= 26.44 & [p = 0.000]
 \end{aligned}$$

(absolute value of t statistics in brackets).

Now consider the distribution of wages for these cohorts later in life. From Table 2, we see that the Gini coefficient in wages is smaller than that of the corresponding test score in all but four cases. In each of these cases the difference is not numerically large, and a comparison of the full Lorenz curves strongly suggests that the Gini coefficients are essentially identical in these four cases.

When looking at wages, we will first focus on the shaded observations in Table 2. These observations capture cohort I in their late thirties (in the mid to late 1980s), and cohort II as

late as possible. In both cohorts the United States had the highest degree of wage dispersion. Australia had the least disperse wages in the first cohort, but it did not participate in the second IME. Of the cohort II participants, Japan had the lowest wage dispersion.

That wage Gini coefficients are never significantly larger than test score coefficients may be spurious given the small number of countries. However, their values overlap a great deal: in some countries G_w is greater than G_s for one or more other countries. Comparing Lorenz curves offers a more complete explanation. Figure 2 graphs the Lorenz curves for the selected (shaded) observations for cohort I, and Figure 3 graphs the curves for cohort II. The whole test score Lorenz curve lies below the whole wage Lorenz curve in each of the first cohort observations, except for some slight overlap in Finland where the curves are nearly identical.

Consider a few specific countries in cohort I: the United Kingdom, the United States, Australia, and Finland. The U.K. had both the most unequal distribution of test scores (as shown by Figure 2) and the largest proportion of students with very low scores. The wage distribution, however, is much more equal. In contrast, test scores in the U.S. are somewhat less disperse than in the U.K. but the gap between wage and test score dispersion is much smaller. In this sense, the U.K. has the potential to have the worst wage distribution within this cohort but in fact ranks third. Australia and Finland make for a similar comparison at the other end of the spectrum. Finland has the most equal distribution of test scores and very few people scoring at the bottom. However, Finland's wage distribution is nearly the same as its score distribution and is more unequal than many other countries including Australia, which like the U.K. has substantially less wage inequality than score inequality.

The story is somewhat different for the second IME cohort shown in Figure 3. Dispersion in test scores appears to have fallen consistently across countries. The one exception, Finland, is notable since it had the least score dispersion in cohort I but is now surpassed in that regard by several countries including the United States. Despite this general shift up in the test score Lorenz curves, the selected wage observations still lay on top of, or inside the score curves. Here the one exception is the U.S., where the wage curve drops slightly below the

score curve at the top of the distribution, leading to a slightly higher Gini coefficient.

Test Scores and Pupil-Teacher Ratios

The IME examination provides a snapshot of the performance of primary schools before students diverge in their studies during high school. Can dispersion in test scores at this age be explained by measurable characteristics of primary schools? UNESCO reports several educational statistics for countries going back to the 1960s. We look at pupil-teacher ratios (PTR) in primary schools, because it gets closest to measuring difference in the average classroom across countries.

$$G_s = .218 - .0019 \times \text{PTR} - .0640 \times \text{CohortII}$$

(5.2) (1.27) (3.49)

$$F_{2,12} = 6.81 \quad [p = 0.0106],$$

The statistically significant cohort effect in the estimated regression suggests that a difference in the testing instrument accounts for the fall in test score dispersion. But it cannot be ruled out that primary schools changed systematically between the 1950s and the 1970s so as to lower the dispersion of student skills across cohorts. Table 1 shows that several countries experience a sharp decline in pupil-teacher ratios over this time period, including the U.S., Japan, the Netherlands, and Finland. Only Sweden had a slight increase in the PTR, and only Finland had an increase in score dispersion. Within cohorts, however, the regression indicates a weak *negative* relationship between PTR and wage dispersion. That is, countries with larger classrooms tend to have *less* score dispersion.

These results might be interpreted as saying that big classrooms are good for all students. However, it is clear that many other factors might lead to this unusual result. Heyneman, Stephen and William Loxley (1983), having access to some characteristics of the individuals in the second IME cohort, report regressions that suggest that family background is more important in developed countries and that different school variables are important in different countries. At the least, the result suggests that international differences in test score distributions are not arbitrary and may reflect differences in school systems.

The Score-Wage Gap

Figures 2 and 3 suggest that the wage and score Lorenz curves move together. Is it then true that the Gini coefficients move together? That is, does overall wage dispersion seem related to overall test score dispersion? Our first regression,

$$G_w = .0488 - .3585 \times G_s - .0008 \times \text{CohortII}$$

(1.61) (1.96) (0.06)

$$F_{2,16} = 4.08 \quad [p = 0.0371],$$

strongly suggests they do. The relationship is stable across the two cohorts, as the coefficient on cohort II is near zero and insignificant. Figure 4 shows the stable relationship between G_s and G_w within and across cohorts. Perhaps the general drop in test score dispersion between the 1962 and 1982 IME cohorts was not an artifact of the test, because wage dispersion seems to have decreased predictably along with the score dispersion.

To some extent, the line in Figure 4 seems to split countries that tend to have decentralized wage setting (the U.S. and the U.K.) and those with centralized wage setting (Sweden, the Netherlands). To explain the residual, we add the country's union density (listed in Table 1) as a measure of wage centralization:

$$G_w = .062 - .342 \times G_s - .0003 \times \text{Union}$$

(3.07) (2.8) (1.42)

$$F_{2,16} = 5.60 \quad [p = 0.0143].$$

After controlling for test score dispersion, greater union density is associated with lower wage dispersion. (The cohort indicator has been dropped since it was insignificant in the first regression.) This is certainly not a surprising finding, but at the least it confirms that our computed Gini coefficients are not spurious, and that controlling for test score dispersion is simply an alternative way to control for interventionist government policy. Our results are consistent with Blau and Kahn's (1996) more extensive analysis of wage centralization and inequality. They find that several measures of wage centralization in nine countries help explain the greater dispersion in male wages in the U.S., particularly at the bottom of the

distribution. We still find a significant role for labor market institutions, but *educational* institutions are also a fundamental explanation for wage dispersion.

We also tried to explain G_w using UNESCO data on secondary and post-secondary school attendance rates and the age at which academic streaming occurs in the IME countries. Whether union density is included or not, these variables are statistically insignificant when G_s is included. Measurable differences in further schooling after age thirteen do not help explain the gap between G_w and G_s across countries.

Three Observations Holding Age Constant

Although there appears to be a steady relationship between wage and score dispersion across the two cohorts, the comparison is based on observations of wages at different ages, when the older cohort is in its late thirties and the younger cohort is in its twenties. For three countries (Japan, the U.K., and the U.S.) we have wage data from the early 1970s when the first cohort was also in its twenties. For these countries Figure 5 compares wage Lorenz curves for the two cohorts at the same age. (The wage Lorenz curves are labeled by cohort. The score curves are not shown. As discussed earlier, $L_s^{II}(z) > L_s^I(z)$, and $L_w(z)$ lies within the corresponding $L_s(z)$.) Wage dispersion at the same age has increased slightly (in Japan stayed nearly constant) across the two cohorts, despite the fifteen year difference and the large shift in test scores. We define $L_w^*(z)$ as the predicted value of the Lorenz curve for cohort II given their test scores and the relationship between $L_w(z)$ and $L_s(z)$ for the earlier cohort. That is, define

$$\lambda(z) \equiv \frac{z - L_w^I(z)}{z - L_s^I(z)}$$

as the proportion of the test-score gap that is made up in wages by the bottom z percent of cohort I in the early 1970s. If we hold $\lambda(z)$ constant and treat the second cohort's score distribution, $L_s^{II}(z)$, as the new distribution of the same skills as cohort I, then the predicted Lorenz curve for cohort II is

$$L_w^*(z) \equiv z - \lambda(z) \left(z - L_s^{II}(z) \right).$$

For all three countries $L_w^*(z)$, the unlabelled thick curve, lies well within both of the actual wage Lorenz curves. That is, conditional on the cohort's distribution of math skills, there is more wage dispersion among the second cohort than for the earlier cohort at the same age. This may suggest a shift in the test instrument itself, but the earlier results suggest that at least some of the differences between the two cohorts' test scores are reflected in wages as of the late 1980s and early 1990s. If so, then the gap between $L_w^*(z)$ and $L_w^{II}(z)$ suggests a change in the relationship between wages and math skills from the early 1970s to the mid 1980s.

To summarize, our two inter-cohort comparisons, one holding calendar time (relatively) constant and one holding cohort age (relatively) constant, suggest two consistent explanations. One is that the drop in score dispersion among the second IME cohort is primarily an artifact of the tests, and the 'excess' wage dispersion in Figure 5 experienced by the second cohort while in their twenties is not excessive. This explanation would depend on the steady relationship in Figure 4 between G_w and G_s being a coincidence in how the exam changed between cohorts. An alternative explanation is that the drop in test score dispersion between the two cohorts was real and the excess wage dispersion in Figure 5 reflects a general increase in skill-related wage dispersion. Typically, skill-biased technological change is suggested by increased residual variance in wages. Our results are conditioned upon a measure of skill dispersion independent of any labor market outcomes, and may be interpreted as more direct evidence for skill-biased technological change than evidence based on wage-data alone.

IV. Conclusion

We have compared the distribution of scores on math test given at age thirteen to the distribution of wages later in life for two cohorts of people in eleven countries. Our results are as follows: wage dispersion is consistently lower than test score dispersion across both countries and time; test score dispersion fell between 1962 and 1982; this fall in dispersion

is at least partly reflected in wage dispersion in the late 1980s; the fall in score dispersion masks a marginally significant negative relationship between dispersion and average class size in primary school; union density helps explain the difference between wage inequality and test score dispersion.

Our results suggest that countries historically differed greatly in the distribution of skills they imparted to students, that these skills help determined wages later in life, and that international differences in pre-market skills help explain international differences in wage dispersion. Our results also suggest a convergence in the distributions of skills over the last thirty years has occurred in several industrialized countries. As a consequence, increases in wage inequality may understate the effect of labor market forces such as government wage policy and technological change because younger people appear to be bringing more equal skills to the market.

References

- Bedard, Kelly (1997), "Educational Streaming, Occupational Choice, and the Distribution of Wages," manuscript, McMaster University.
- Blau, and Kahn (1996), "International Differences in Male Wage Inequality: Institutions Versus Market Forces," *Journal of Political Economy* 104(4), 791-836.
- Card, David and Alan Krueger (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100(1), 1-40.
- Cramer, John and George Cowen Brown (1965), *Contemporary Education: A Comparative Study of National Systems*, New York: Harcourt, Brace, and World Inc.
- Davis, Steven (1992), "Cross-Country Patterns in Relative Wages," in *NBER Macroeconomics Annual*, Blanchard, Olivier and Stanley Fischer (eds), Cambridge: MIT Press, 239-92.
- Freudenthal, Hans (1975), "Pupils' Achievement Internationally Compared - the IEA," *Educational Studies in Mathematics*, 6.
- Gottschalk and Joyce (1992), "Changes in Earnings Inequality: An International Perspective," working paper 232, Boston University, December.
- Heyneman, Stephen and William Loxley (1983), "The Effect of Primary School Quality on Academic Achievement across Twenty-nine High- and Low-Income Countries," *American Journal of Sociology* 88(6), 1162-1194.
- Husen, Torsten (1967), *International Study of Achievement in Mathematics, a Comparison of Twelve Countries*, Stockholm: Almqvist and Wiksell, Volumes 1 and 2.
- Keeves, John (1998), "Cross-National Comparisons in Educational Achievement: the Role of the International Association for the Evaluation of Educational Achievement (IEA)," in S.P. Heneman and I. Fagerlind (eds.), *University Examinations and Standardized Tests: Principles, Experience, and Policy Options*, World Bank Technical Paper No. 78.

OECD (1989-1990), *Education in OECD Countries*

OECD (1981), *OECD Studies in Taxation*

Paglin, Morton and Anthony Rufolo (1990), "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences," *Journal of Labor Economics* 8(1), 123-144.

Robitaille, David (1990), "Canadian Participation in the Second International Mathematics Study," working paper, Economic Council of Canada No. 6.

Roy, Andrew (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135-146.

Traxler, Franz (forthcoming), "Collective Bargaining and Industrial Change: A Case of Disorganization" *European Sociological Review*.

UNESCO (1964, 1986, 1987), *Yearbook*.

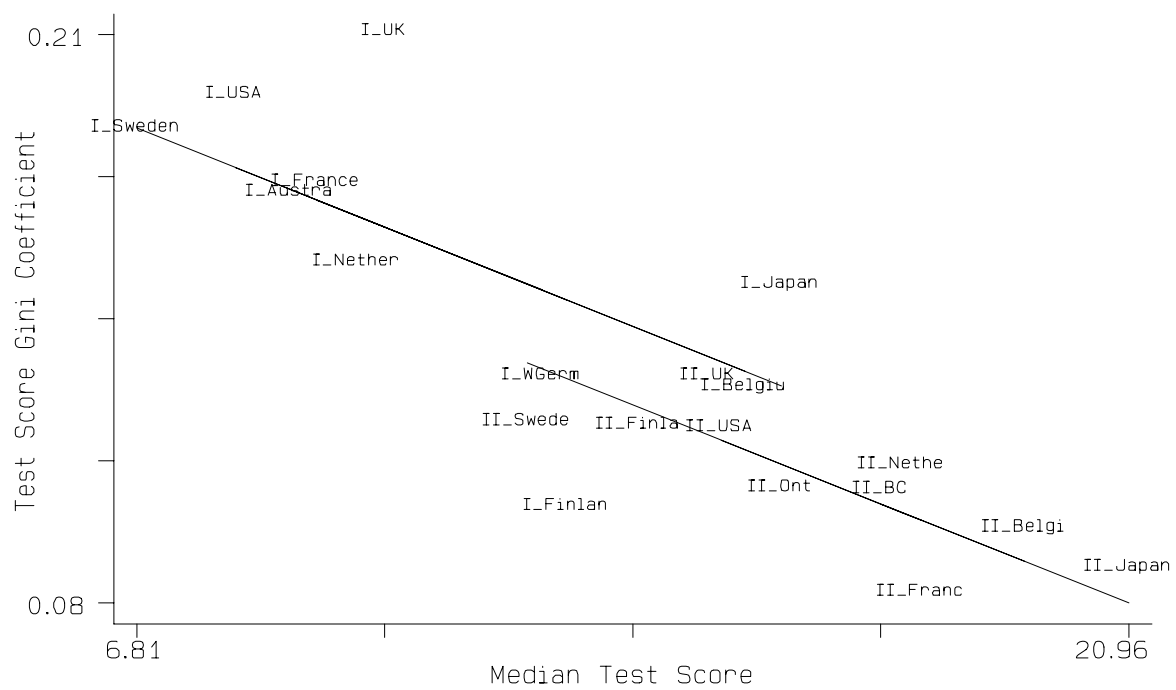


Figure 1. Median Test Scores and Test Score Dispersion

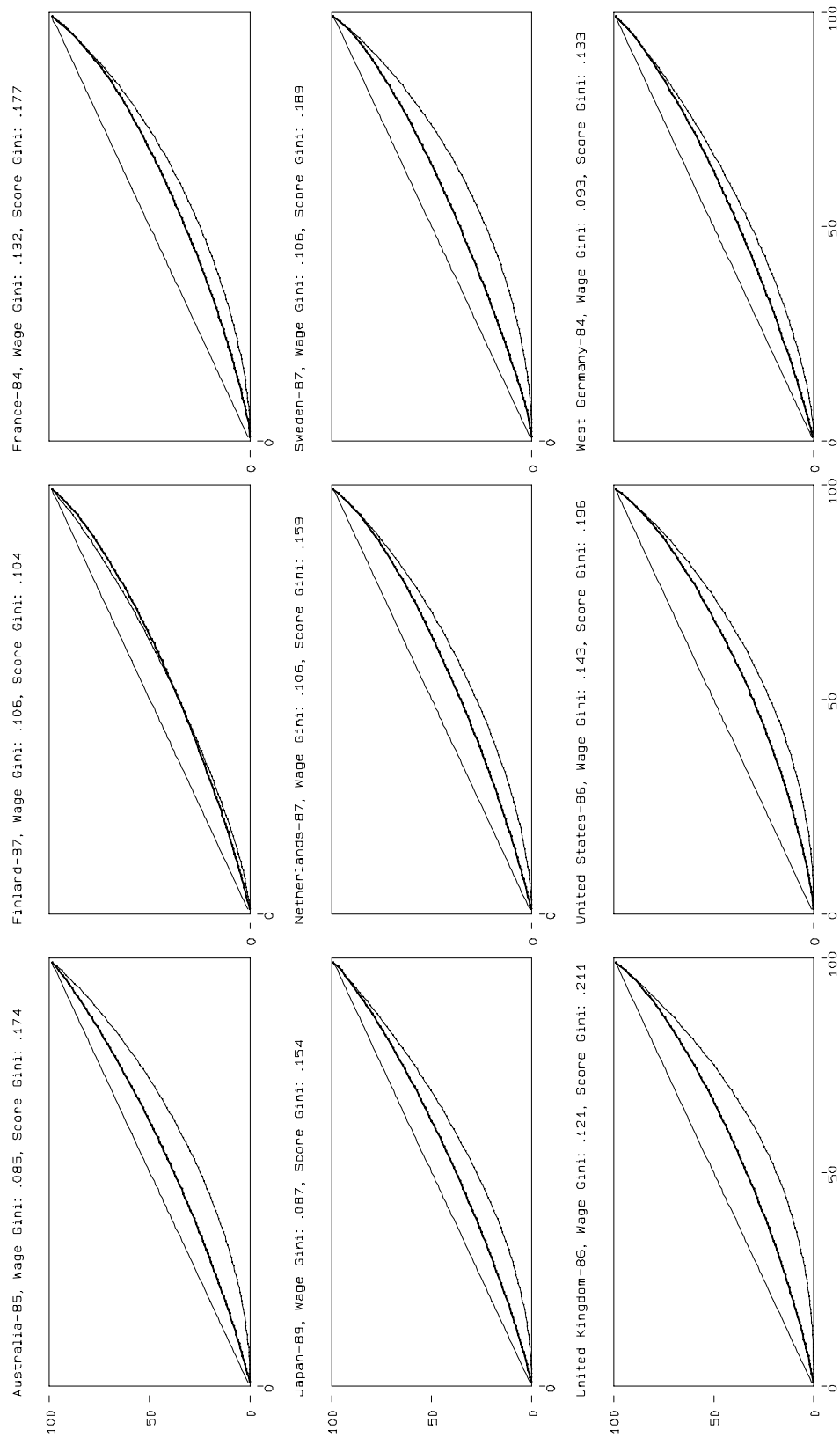


Figure 2. Wage (thick) and Score Lorenz Curves, Cohort I

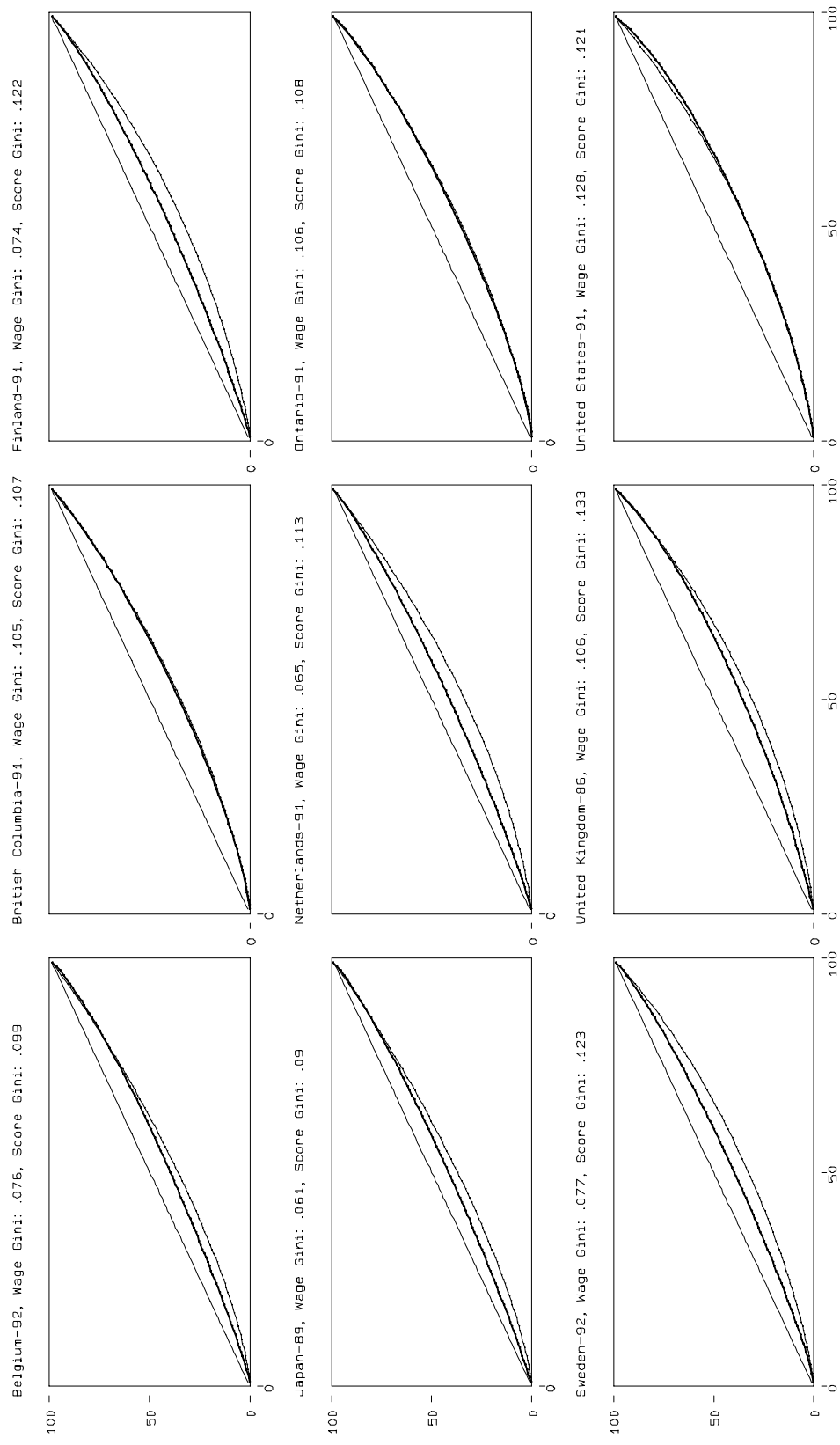


Figure 3. Wage (thick) and Score Lorenz Curves, Cohort II



Figure 4. Wage and Test Score Dispersion

Table 1.
Sources of Wage Data, Sample Sizes, Pupil-Teacher Ratios, Union Density

Country	Original Data Source for Wage Data	Full Time			Coh.	Wage Obs 1		Wage Obs 2		Wage Obs 3		Wage Obs 4		Wage Obs 5		P.T.R.	Union
		A	B	C		N	Year	N	Year	N	Year	N	Year	N	Year		
1 B. C.	<i>Census of Canada Public Use Files</i>				II	598	86	1410	91							.	35.8
2 Japan	<i>Census of Wages</i>			*	I	1904	69	2741	74	1775	79	1679	83	23243	89	25	25.4
					II	4650	89									25	25.4
3 Ontario	<i>Census of Canada Public Use Files</i>				II	3355	86	6254	91							25	25.4
4 U. K.	<i>Family Expenditure Survey</i>	*		*	I	526	74	444	86							25	39.1
					II	416	86									18	39.1
5 Australia	<i>Income and Housing Survey</i>	*	*	*	I	1184	81	561	85	882	89					30	40.4
6 Belgium	<i>Living Conditions of Households</i>	*	*		I	520	85	308	88	250	92					.	51.4
					II	390	85	289	88	284	92					.	51.4
7 Finland	<i>Survey of Income Distribution</i>	*	*		I	923	87	851	91							24	72
					II	645	87	541	91							15	72
8 France	<i>French Income Survey of Taxes</i>	*			I	773	79	896	84							29	9.8
					II	275	84									21	9.8
9 W. Germ.	<i>German Panel Survey: Wave 2</i>	*	*	*	I	3567	78	192	81	412	84					29	32.9
10 Nether.	<i>Survey of Income and Program Use</i>	*	*		I	454	83	367	87	313	91					33	25.5
					II	307	87	365	91							20	25.5
11 Sweden	<i>Swedish Income Distribution Survey</i>	*		*	I	728	75	620	81	749	87	871	92			15	82.5
					II	415	87	520	92							16	82.5
12 U.S.A.	<i>Current Population Survey</i>	*	*	*	I	416	69	858	74	973	79	684	86	891	91	24	15.6
					II	654	86	1064	91							21	15.6

Notes

2 Our Source: Handbook of Labor Statistics; Sample Sizes are in thousands

6 Earnings converted to pre-tax using schedules in OECD (1987)

A 30 hours or more of work per week.

B 46 or more weeks per year.

C Full time employee

All samples are also restricted to: male, not self-employed,

Source for Pupil Teacher Ratio: UNESCO (1964, 1986, 1987)

Source for Union Density: Traxler (forthcoming)

Table 2.
Wages Later in Life for the Two IME Cohorts

IME Country	IME 13-year-olds				Distribution of Wages Later in Life for IME Cohort									
	IME Coh.	Sam. Size	Median Score	Score Gini	Obs. 1		Obs. 2		Obs. 3		Obs. 4		Obs. 5	
					A	Gini	A	Gini	A	Gini	A	Gini	A	Gini
1 B. C.	II	2167	17.4	0.107	24	0.123	29	0.105						
2 Japan	I	2050	16.0	0.154	22	0.077	27	0.066	32	0.065	37	0.071	42	0.087
	II*	8091	21.0	0.090	27	0.061								
3 Ontario	II	4666	16.0	0.108	24	0.121	27	0.106						
4 U. K.	I*	3552	10.4	0.211	27	0.083	39	0.121						
	II*	2483	15.0	0.133	24	0.106								
5 Australia	I*	3078	9.0	0.174	34	0.092	38	0.085	42	0.110				
6 Belgium	I	2645	15.5	0.131	38	0.097	41	0.094	45	0.100				
	II	2724	19.5	0.099	23	0.076	26	0.072	30	0.076				
7 Finland	I*	841	12.9	0.104	40	0.106	44	0.104						
	II	4382	14.0	0.122	25	0.079	29	0.074						
8 France	I	3449	9.4	0.177	32	0.112	37	0.132						
	II*	2483	18.0	0.084	22	0.093								
9 W. Germ.	I	4475	12.6	0.133	33	0.075	34	0.078	37	0.093				
10 Nether.	I	1433	10.0	0.159	36	0.088	40	0.106	44	0.105				
	II	5418	17.7	0.113	25	0.065	29	0.065						
11 Sweden	I	2828	6.8	0.189	28	0.069	34	0.081	40	0.106	45	0.105		
	II	3451	12.4	0.123	25	0.076	30	0.077						
12 U.S.A.	I*	6544	8.2	0.196	22	0.113	27	0.101	32	0.114	39	0.143	44	0.138
	II*	6648	15.1	0.121	24	0.136	29	0.128						

Wages from four-year age category centered on age A (see Table 1 for more details).

Bold = Wage coefficient > Score coefficient;

* = row contains a score or selected wage coefficient that is a minimum or maximum.

Shading indicates the observation chosen to represent the country and IME cohort.